

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Horvitz, Robert
Yuan, Junying
Shaham, Shai

(ii) TITLE OF THE INVENTION: RELATEDNESS OF HUMAN INTERLEUKIN-1
BETA CONVERTASE GENE TO A C. ELEGANS CELL DEATH
GENE, INHIBITORY PORTIONS OF THESE GENES AND...

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Clark & Elbing LLP
(B) STREET: 176 Federal Street
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 21-MAY-1995
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/394,189
(B) FILING DATE: 24-FEB-1995

(A) APPLICATION NUMBER: 08/282,211
(B) FILING DATE: 12-JUL-1994

(A) APPLICATION NUMBER: 07/984,182
(B) FILING DATE: 20-NOV-1992

(A) APPLICATION NUMBER: 07/897,788
(B) FILING DATE: 12-JUN-1992

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Bieker-Brady, Kristina
(B) REGISTRATION NUMBER: 39,109

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617-428-0200

(B) TELEFAX: 617-428-7045

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7653 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

15 AGATCTGAAA TAAGGTGATA AATTAATAAAA TTAAGTGTAT TTCTGAGGAA ATTTGACTGT 60
TTTAGCACAA TTAATCTTGT TTCAGAAAAA AAGTCCAGTT TTCTAGATTT TTCCGTCTTA 120
TTGTCGAATT AATATCCCTA TTATCACTTT TTCATGCTCA TCCTCGAGCG GCACGTCCTC 180
AAAGAAATTGT GAGAGCAAAC GCGCTCCCAT TGACCTCCAC ACTCAGCCGC CAAAACAAAC 240
GTTCTGAACAT TCGTGTGTTG TGCTCCTTTT CCGTTATCTT GCAGTCATCT TTTGTCGTTT 300
TTTTCTTTGT TCTTTTTTGT GAACGTGTTG CTAAGCAATT ATTACATCAA TTGAAGAAAA 360
20 GGCTCGCCGA TTTATTGTTG CCAGAAAGAT TCTGAGATTC TCGAAGTCGA TTTTATAATA 420
TTTAACCTTG GTTTTTGTCAT TGTTTCGTTT AAAAAACCA CTGTTTATGT GAAAAACGAT 480
TAGTTTACTA ATAAACTAC TTTTAAACCT TTACCTTTAC CTCACCGCTC CGTGTTTCATG 540
GCTCATAGAT TTTGATACT CAAATCCAAA AATAAATTTA CGAGGGCAAT TAATGTGAAA 600
CAAAAACAAT CCTAAGATTT CCACATGTTT GACCTCTCCG GCACCTTCTT CCTTAGCCCC 660
25 ACCACTCCAT CACCTCTTTG GCGGTGTTCT TCGAAACCCA CTTAGGAAAG CAGTGTGTAT 720
CTCATTTGGT ATGCTCTTTT CGATTTTATA GCTCTTTGTC GCAATTTCAA TGCTTTAAAC 780
AATCCAAATC GCATTATATT TGTGCATGGA GGCAAATGAC GGGGTGGA TCTTAGATGA 840
GATCAGGAGC TTTCAGGGTA AACGCCCGGT TCATTTTGTA CCACATTTCA TCATTTTCCT 900
GTCGTCTTG GTATCCTCAA CTTGTCCCGG TTTTGTTTTC GGTACACTCT TCCGTGATGC 960
30 CACCTGTCTC CGTCTCAATT ATCGTTTAGA AATGTGAACT GTCCAGATGG GTGACTCATA 1020
TTGCTGCTGC TACAATCCAC TTTCTTTTCT CATCGGCAGT CTTACGAGCC CATCATAAAC 1080
TTTTTTTTTCC GCGAAATTTG CAATAAACCG GCCAAAAACT TTCTCCAAAT TGTTACGCAA 1140
TATATACAAT CCATAAGAAT ATCTTCTCAA TGTTTATGAT TTCTTCGCAG CACTTTCTCT 1200
TCGTGTGCTA ACATCTTATT TTTATAATAT TTCCGCTAAA ATTCCGATTT TTGAGTATTA 1260
35 ATTTATCGTA AAATTATCAT AATAGCACCG AAACTACTA AAAATGGTAA AAGCTCCTTT 1320
TAAATCGGCT CGACATTATC GTATTAAAGGA ATCACAATAA TCTGAGAATG CGTACTGCGC 1380
AACATATTTG ACGGCAAAAT ATCTCGTAGC GAAAACTACA GTAATTCCTT AAATGACTAC 1440
TGTAGCGCTT GTGTCGATTT ACGGGCTCAA TTTTGTAAAA TAATTTTTTT TTTTCAATTT 1500
TGATAACCCG TAAATCGTCA CAACGCTACA GTAGTCATTT AAAGGATTAC TGTAGTTCTA 1560
40 GCTACGAGAT ATTTTGC GCGC CCAAATATGA CTGTAATACG CATTCTCTGA ATTTTGTGTT 1620
TCCGTAATAA TTTTACAAGA TTTTGGCATT CCACCTTTAAA GGCGCACAGG ATTTATTCCA 1680
ATGGGTCTCG GCACGCAAAA AGTTTGATAG ACTTTTAAAT TCTCCTTGCA TTTTAAATTC 1740
AATTACTAAA ATTTTTCGTA ATTTTCTGT TAAAATTTTT AAAATCAGTT TTCTAATATT 1800
TTCCAGGCTG ACAAACAGAA ACAAACAC AACAAACATT TTAATAATCA GTTTTCAAAT 1860

	TAAAAATAAC	GATTTCTCAT	TGAAAATTGT	GTTTTATGTT	TGCGAAAATA	AAAGAGAACT	1920
	GATTCAAAAC	AATTTTAAACA	AAAAAAAACC	CCAAAATTCG	CCAGAAATCA	AGATAAAAAA	1980
	TTCAAGAGGG	TCAAAATTTT	CCGATTTTAC	TGACTTTTAC	CTTTTTTTTC	GTAGTTCAGT	2040
	GCAGTTGTTG	GAGTTTTTTGA	CGAAAACTAG	GAAAAAAATC	GATAAAAATT	ACTCAAATCG	2100
5	AGCTGAATTT	TGAGGACAAT	GTTTAAAAAA	AAACACTATT	TTTCCAATAA	TTTCACTCAT	2160
	TTTCAGACTA	AATCGAAAAT	CAAATCGTAC	TCTGACTACG	GGTCAGTAGA	GAGGTCAACC	2220
	ATCAGCCGAA	GATGATGCGT	CAAGATAGAA	GGAGCTTGCT	AGAGAGGAAC	ATTATGATGT	2280
	TCTCTAGTCA	TCTAAAAGTC	GATGAAATTY	TCGAAGTTCT	CATCGCAAAA	CAAGTGTTGA	2340
10	ATAGTGATAA	TGGAGATATG	ATTAATGTGA	GTTTTTAATC	GAATAATAAT	TTTAAAAAAA	2400
	AATTGATAAT	ATAAAGAATA	TTTTTGCAGT	CATGTGGAAC	GGTTCGCGAG	AAGAGACGGG	2460
	AGATCGTGAA	AGCAGTGCAA	CGACGGRGAG	ATGTGGCGTT	CGACGCGTTT	TATGATGCTC	2520
	TTGCTCTAC	GGGACACGAA	GGACTTGCTG	AAGTTCCTGA	ACCTCTCGCC	AGATCGTAGG	2580
	TTTTTAAAGT	TCGGCGCAAA	AGCAAGGGTC	TCACGGAAAA	AAGAGGCGGA	TCGTAATTTT	2640
	GCAACCCACC	GGCACGGTTT	TTTCTCCGA	AAATCGGAAA	TTATGCACTT	TCCCAAATAT	2700
15	TTGAAGTGAA	ATATATTTTA	TTTACTGAAA	GCTCGAGTGA	TTATTTATTT	TTTAACACTA	2760
	ATTTTCGTGG	CGCAAAAGGC	CATTTTGTAG	ATTTGCCGAA	AATACTTGTC	ACACACACAC	2820
	ACACACATCT	CCTTCAAATA	TCCCTTTTTT	CAGTGTGAC	TCGAATGCTG	TCGAATTCGA	2880
	GTGTCCAATG	TCACCGGCAA	GCCATCGTCG	GAGCCGCGCA	TTGAGCCCCG	CCGGCTACAC	2940
	TTCACCGACC	CGAGTTTACC	GTGACAGCGT	CTCTTCAGTG	TCATCATTCA	CTTCTTATCA	3000
20	GGATATCTAC	TCAAGAGCAA	GATCTCGTTC	TCGATCGCGT	GCACTTCATT	CATCGGATCG	3060
	ACACAATTAT	TCATCTCCTC	CAGTCAACGC	ATTTCCCAGC	CAACCTTGTA	TGTTGATGCG	3120
	AACACTAAAT	TCTGAGAATG	CGCATTACTC	AACATATTTG	ACGCGCAAAT	ATCTCGTAGC	3180
	GAAAAATACA	GTAACCCTTT	AAATGACTAT	TGTAGTGTG	ATTTACGGGC	TCGATTTTTCG	3240
	AAACGAATAT	ATGCTCGAAT	TGTGACAACG	AATTTTAAAT	TGTCATTTTT	GTGTTTTCTT	3300
25	TTGATATTTT	TGATCAATTA	ATAAATTATT	TCCGTAAACA	GACACCAGCG	CTACAGTACT	3360
	CTTTTAAAGA	GTTACAGTAG	TTTTTCGCTT	AAGATATTTT	GAAAAGAATT	TTAAACATTT	3420
	TGAAAAAAA	TCATCTAACA	TGTGCCAAAA	CGCTTTTTTT	AAGTTTCGCA	GATTTTTTGA	3480
	TTTTTTTTCAT	TCAAGATATG	CTTATTAACA	CATATAATTA	TCATTAATGT	GAATTTCTTG	3540
	TAGAAATTTT	GGGCTTTTTC	TTCTAGTATG	CTCTACTTTT	GAAATTGCTC	AACGAAAAAA	3600
30	TCATGTGGTT	TGTTTCATATG	AATGACGAAA	AATAGCAATT	TTTTATATAT	TTTCCCTAT	3660
	TCATGTTGTG	CAGAAAAATA	GTAAAAAAGC	GCATGCATTT	TTTCGACATTT	TTTACATCGA	3720
	ACGACAGCTC	ACTTCACATG	CTGAAGACGA	GAGACGCGGA	GAAATACCAC	ACATCTTTCT	3780
	GCGTCTCTCG	TCTTCAGCAT	GTGAAATGGG	ATCTCGGTCT	ATGTAAAAAA	ATGTGGAATA	3840
	ATGTAAAAAA	TGCATGCGTT	TTTTTACACT	TTTCTGCACA	AATGAATAGG	GGGAAAAATG	3900
35	ATTAAAATAC	ATTTTTTGTA	TTTTTCAACA	TCACATGATT	AACCCCATTA	TTTTTTCGTT	3960
	GAGCAACTTA	AAAAGTAGAG	AATATTAGAG	CGAAAACCAA	AATTCTTCA	AGATATTACC	4020
	TTTATTGATA	ATTATAGATG	TTAATAAGCA	TATCTTGAAT	GAAAGTCAGC	AAAAATATGT	4080
	GCGAAACACC	TGAAAAAAAT	CAAAAATTCT	GCGAAAATTG	AAAAAATGCA	TTAAAATACA	4140
	TTTTTGCATT	TTTCTACATC	ACATGAATGT	AGAAAATTAA	AAGGGAAATC	AAAATTCTTA	4200
40	GAGGATATAA	TTGAATGAAA	CATTGCGAAA	TTAAAATGTG	CGAAACGTCA	AAAAAGAGGA	4260
	AATTTGGGTA	TCAAAATCGA	TCCTAAAACC	AACACATTTT	AGCATCCGCC	AACTCTTCAT	4320
	TCACCGGATG	CTCTTCTCTC	GGATACAGTT	CAAGTCGTAA	TCGCTCATTC	AGCAAAGCTT	4380
	CTGGACCAAC	TCAATACATA	TTCCATGAAG	AGGATATGAA	CTTTGTTCGAT	GCACCAACCA	4440
	TAAGCCGTGT	TTTCGACGAG	AAAACCATGT	ACAGAAACTT	CTCGAGTCCT	CGTGGAATGT	4500
45	GCCTCATCAT	AAATAATGAA	CACTTTGAGC	AGATGCCAAC	ACGGAATGGT	ACCAAGGCCG	4560
	ACAAGGACAA	TCTTACCAAT	TTGTTTCAGAT	GCATGGGCTA	TACGGTTATT	TGCAAGGACA	4620
	ATCTGACGGG	AAGGGTACGG	CGAAATTATA	TTACCCAAAC	GCGAAATTTG	CCATTTTTCG	4680
	CCGAAAATGT	GGCGCCCGGT	CTCGACACGA	CAATTTGTGT	TAAATGCAAA	AATGTATAAT	4740
	TTTGCAAAAA	ACAAAATTTT	GAACCTCCGC	GAAAATGATT	TACCTAGTTT	CGAAATTTTC	4800
50	GTTTTTTTCCG	GCTACATTAT	GTGTTTTTTT	TTAGTTTTTT	TATAATATTT	GATGTAAAAA	4860
	ACCGTTTGTA	AATTTTCAGA	CAATTTTCCG	CATACAAAAC	TTGATAGCAC	GAAATCAATT	4920

	TTCTGAATTT	TCAAAATTAT	CCAAAAATGC	ACAAATTAAA	ATTTGTGAAA	ATTGGCAAAC	4980
	GGTGTTTTCAA	TATGAAATGT	ATTTTTTAAAA	ACTTTAAAAA	CCACTCCGGA	AAAAGCAATAA	5040
	AAATCAAAAC	AACGTACAA	TTCAAATTCA	AAAGTTATTC	ATCCGATTTG	TTTATTTTTTG	5100
	CAAAATTTGA	AAAAATCATG	AAGGATTTAG	AAAAGTTTTA	TAACATTTTTT	TCTAGATTTTT	5160
5	TCAAAATTTT	TTTTTAACAA	TCGAGAAAAA	GAGAATGAAA	AATCGATTTT	AAAAATATCC	5220
	ACAGCTTCGA	GAGTTTGAAA	TTACAGTACT	CCTTAAAGGC	GCACACCCCA	TTTGCATTGG	5280
	ACCAAAAATT	TGTCGTGTCTG	AGACCAGGTA	CCGTAGTTTT	TGTCGCAAAA	ATTGCACCAT	5340
	TGGACAATAA	ACCTTCCCTAA	TCACCAAAAA	GTAAAAATTGA	AATCTTCGAA	AAGCCAAAAA	5400
	ATTCAAAAAA	AAAGTCGAAT	TTCGATTTTT	TTTTTGGTTT	TTTGGTCCCA	AAAACCAAAA	5460
10	AAATCAATTT	TC TGCAAAA	ACCAAAAAGA	AACCCGAAAA	AAATTTCCCAG	CCTTGTTCCT	5520
	AATGTAACT	GATATTTAAT	TTCCAGGGAA	TGCTCCTGAC	AAATCGAGAC	TTTGCCAAAC	5580
	ACGAATCACA	CGGAGATTCT	GCGATACTCG	TGATTCTATC	ACACGGAGAA	GAGAATGTGA	5640
	TTATTGGAGT	TGATGATATA	CCGATTAGTA	CACACGAGAT	ATATGATCTT	CTCAACGCGG	5700
	CAAATGCTCC	CCGTCTGGCG	AATAAGCCGA	AAATCGTTTT	TGTGCAGGCT	TGTCGARGCG	5760
15	GTTCTGTTTT	TATTTTAAAT	TTAATATAAA	TATTTTAAAT	AAATTCATTT	TCAGAACGTC	5820
	GTGACAATGG	ATTCCCAGTC	TTGGATTCTG	TCGACGGAGT	TCCTGCAATTT	CTTCGTCTGT	5880
	GATGGGACAA	TCGAGACGGG	CCATTGTTCA	ATTTTCTTGG	ATGTGTGCGG	CCGCAAGTTY	5940
	AGGTTGCAAT	TTAATTTCTT	GAATGAGAAT	ATTCCTTCAA	AAAACTTAAA	ATAGATTTTT	6000
	ATTCCAGAAA	GTCCCAGTCG	AAAAATTGCG	ATATAATTAC	GAAATTTGTG	ATAAAATGAC	6060
20	AAACCAATCA	GCATCGTCGA	TCTCCGCCCA	CTTCATCGGA	TTGGTTTTGAA	AGTGGGCGGA	6120
	GTGAATTGCT	GATTGGTCGC	AGTTTTTCAGT	TTAGAGGGAA	TTTTAAAAATC	GCCTTTTTCGA	6180
	AAATTAAAAA	TTGATTTTTT	CAATTTTTTTC	GAAAAATATT	CCGATTATTT	TATATTCTTT	6240
	GGAGCGAAAG	CCCCGTCTTG	TAAACATTTT	TAAATGATAA	TTAATAAAAT	TTTGCARCAA	6300
	GTGTGGAGAA	AGAAGCCGAG	CYAAGCTGAC	ATTTCTGATT	GRTACGCAAC	GACAGCTCAA	6360
25	TATGTTTTCGT	GGAGAAACAG	TGCTCGTGGA	TCATGGTTTCA	TTCAAGCCGT	CTGTGAAGTG	6420
	TTCTCGACAC	ACGYAAAGGA	TATGGATGTT	GTTGAGCTGC	TGACTGAAGT	CAATAAGAAG	6480
	GTCGYTTGTG	GATTTTCAGAC	ATCACAGGGA	TCGAATATTT	TGAAACAGAT	GCCARAGGTA	6540
	CTTGAAACAA	ACAATGCATG	TCTAACTTTT	AAGGACACAG	AAAAATAGGC	AGAGGCTCCT	6600
	TTTGCAAGCC	TGCGCGCGCT	CAACCTAGAA	TTTTAGTTTT	TAGCTAAAAT	GATTGATTTT	6660
30	GAATATTTTA	TGCTAATTTT	TTTGCGTTAA	ATTTTGAAAT	AGTCACTATT	TATCGGGTTT	6720
	CCAGTAAAAA	ATGTTTATTA	GCCATTGGAT	TTTACTGAAA	ACGAAAATTT	GTAGTTTTTTC	6780
	AACGAAATTT	ATCGATTTTT	AAATGTAAAA	AAAAATAGCG	AAAATTACAT	CAACCATCAA	6840
	GCATTTAAGC	CAAAATTGTT	AACCTCATTTA	AAAATTAATT	CAAAGTTGTC	CACGAGTATT	6900
	ACACGGTTGG	CGCGCGGCAA	GTTTGCAAAA	CGACGCTCCG	CCTCTTTTTTC	TGTGCGGCTT	6960
35	GAAAACAAGG	GATCGGTTTA	GATTTTTCCC	CAAAATTTAA	ATTAAATTTT	AGATGACATY	7020
	CCGCCTGCTC	AAAAAGTTCT	ACTTTTGGCC	GGAAGCACGA	AACTCTGCCG	TCTAAAATTC	7080
	ACTCGTGATT	CATTGCCCAA	TTGATAATTG	TCTGTATCTT	CTCCCCAGT	TCTCTTTCGC	7140
	CCAATTAGTT	TAAACCATG	TGTATATTGT	TATCCTATAC	TCATTTCACT	TTATCATTTCT	7200
	ATCATTTCTC	TTCCCATTTT	CACACATTTT	CATTTCTCTA	CGATAATCTA	AAATTATGAC	7260
40	GTTTGTGTCT	CGAACGCATA	ATAATTTTAA	TAACTCGTTT	TGAATTTGAT	TAGTTGTTGT	7320
	GCCCAGTATA	TATGTATGTA	CTATGCTTCT	ATCAACAAAA	TAGTTTCATA	GATCATCACC	7380
	CCAACCCAC	CAACCTACCG	TACCATATTC	ATTTTTGCCG	GGAATCAATT	TCGATTAATT	7440
	TTAACCTATT	TTTTCGCCAC	AAAAAATCTA	ATATTTGAAT	TAACGAATAG	CATTCCCATC	7500
	TCTCCCGTGC	CGGAATGCCT	CCCGGCCTTT	TAAAGTTCGG	AACATTTGGC	AATTATGTAT	7560
45	AAATTTGTAG	GTCCCCCCCC	TCATTTCCCG	CCCATCATCT	CAAATTGCAT	TCTTTTTTTCG	7620
	CCGTGATATC	CCGATTCTGG	TCAGCAAAGA	TCT			7653

(2) INFORMATION FOR SEO ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 503 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

	Met	Met	Arg	Gln	Asp	Arg	Arg	Ser	Leu	Leu	Glu	Arg	Asn	Ile	Met	Met
	1				5					10					15	
10	Phe	Ser	Ser	His	Leu	Lys	Val	Asp	Glu	Ile	Xaa	Glu	Val	Leu	Ile	Ala
				20				25						30		
	Lys	Gln	Val	Leu	Asn	Ser	Asp	Asn	Gly	Asp	Met	Ile	Asn	Ser	Cys	Gly
			35					40					45			
	Thr	Val	Arg	Glu	Lys	Arg	Arg	Glu	Ile	Val	Lys	Ala	Val	Gln	Arg	Arg
		50					55					60				
15	Xaa	Asp	Val	Ala	Phe	Asp	Ala	Phe	Tyr	Asp	Ala	Leu	Arg	Ser	Thr	Gly
	65					70					75					80
	His	Glu	Gly	Leu	Ala	Glu	Val	Leu	Glu	Pro	Leu	Ala	Arg	Ser	Val	Asp
					85					90					95	
	Ser	Asn	Ala	Val	Glu	Phe	Glu	Cys	Pro	Met	Ser	Pro	Ala	Ser	His	Arg
20				100					105					110		
	Arg	Ser	Arg	Ala	Leu	Ser	Pro	Ala	Gly	Tyr	Thr	Ser	Pro	Thr	Arg	Val
				115				120						125		
	His	Arg	Asp	Ser	Val	Ser	Ser	Val	Ser	Ser	Phe	Thr	Ser	Tyr	Gln	Asp
		130				135						140				
25	Ile	Tyr	Ser	Arg	Ala	Arg	Ser	Arg	Ser	Arg	Ser	Arg	Ala	Leu	His	Ser
	145					150					155					160
	Ser	Asp	Arg	His	Asn	Tyr	Ser	Ser	Pro	Pro	Val	Asn	Ala	Phe	Pro	Ser
					165					170					175	
	Gln	Pro	Ser	Ser	Ala	Asn	Ser	Ser	Phe	Thr	Gly	Cys	Ser	Ser	Leu	Gly
30				180					185					190		
	Tyr	Ser	Ser	Ser	Arg	Asn	Arg	Ser	Phe	Ser	Lys	Ala	Ser	Gly	Pro	Thr
				195				200					205			
	Gln	Tyr	Ile	Phe	His	Glu	Glu	Asp	Met	Asn	Phe	Val	Asp	Ala	Pro	Thr
		210					215					220				
35	Ile	Ser	Arg	Val	Phe	Asp	Glu	Lys	Thr	Met	Tyr	Arg	Asn	Phe	Ser	Ser
	225					230					235					240
	Pro	Arg	Gly	Met	Cys	Leu	Ile	Ile	Asn	Asn	Glu	His	Phe	Glu	Gln	Met
					245					250					255	
	Pro	Thr	Arg	Asn	Gly	Thr	Lys	Ala	Asp	Lys	Asp	Asn	Leu	Thr	Asn	Leu
40				260					265					270		
	Phe	Arg	Cys	Met	Gly	Tyr	Thr	Val	Ile	Cys	Lys	Asp	Asn	Leu	Thr	Gly
				275				280					285			
	Arg	Gly	Met	Leu	Leu	Thr	Ile	Arg	Asp	Phe	Ala	Lys	His	Glu	Ser	His
		290					295					300				
45	Gly	Asp	Ser	Ala	Ile	Leu	Val	Ile	Leu	Ser	His	Gly	Glu	Glu	Asn	Val
	305					310					315					320
	Ile	Ile	Gly	Val	Asp	Asp	Ile	Pro	Ile	Ser	Thr	His	Glu	Ile	Tyr	Asp
					325					330					335	

	Glu	Leu	Leu	Gln	Thr	Arg	Val	Leu	Asn	Lys	Glu	Glu	Met	Glu	Lys	Val	
			30					35					40				
5	AAA	CGT	GAA	AAT	GCT	ACA	GTT	ATG	GAT	AAG	ACC	CGA	GCT	TTG	ATT	GAC	194
	Lys	Arg	Glu	Asn	Ala	Thr	Val	Met	Asp	Lys	Thr	Arg	Ala	Leu	Ile	Asp	
		45					50				55						
	TCC	GTT	ATT	CCG	AAA	GGG	GCA	CAG	GCA	TGC	CAA	ATT	TGC	ATC	ACA	TAC	242
	Ser	Val	Ile	Pro	Lys	Gly	Ala	Gln	Ala	Cys	Gln	Ile	Cys	Ile	Thr	Tyr	
	60					65				70					75		
10	ATT	TGT	GAA	GAA	GAC	AGT	TAC	CTG	GCA	GGG	ACG	CTG	GGA	CTC	TCA	GCA	290
	Ile	Cys	Glu	Glu	Asp	Ser	Tyr	Leu	Ala	Gly	Thr	Leu	Gly	Leu	Ser	Ala	
					80					85					90		
	GAT	CAA	ACA	TCT	GGA	AAT	TAC	CTT	AAT	ATG	CAA	GAC	TCT	CAA	GGA	GTA	338
	Asp	Gln	Thr	Ser	Gly	Asn	Tyr	Leu	Asn	Met	Gln	Asp	Ser	Gln	Gly	Val	
				95					100					105			
15	CTT	TCT	TCC	TTT	CCA	GCT	CCT	CAG	GCA	GTG	CAG	GAC	AAC	CCA	GCT	ATG	386
	Leu	Ser	Ser	Phe	Pro	Ala	Pro	Gln	Ala	Val	Gln	Asp	Asn	Pro	Ala	Met	
				110				115					120				
20	CCC	ACA	TCC	TCA	GGC	TCA	GAA	GGG	AAT	GTC	AAG	CTT	TGC	TCC	CTA	GAA	434
	Pro	Thr	Ser	Ser	Gly	Ser	Glu	Gly	Asn	Val	Lys	Leu	Cys	Ser	Leu	Glu	
		125					130				135						
	GAA	GCT	CAA	AGG	ATA	TGG	AAA	CAA	AAG	TCG	GCA	GAG	ATT	TAT	CCA	ATA	482
	Glu	Ala	Gln	Arg	Ile	Trp	Lys	Gln	Lys	Ser	Ala	Glu	Ile	Tyr	Pro	Ile	
	140					145				150						155	
25	ATG	GAC	AAG	TCA	AGC	CGC	ACA	CGT	CTT	GCT	CTC	ATT	ATC	TGC	AAT	GAA	530
	Met	Asp	Lys	Ser	Ser	Arg	Thr	Arg	Leu	Ala	Leu	Ile	Ile	Cys	Asn	Glu	
					160					165					170		
	GAA	TTT	GAC	AGT	ATT	CCT	AGA	AGA	ACT	GGA	GCT	GAG	GTT	GAC	ATC	ACA	578
	Glu	Phe	Asp	Ser	Ile	Pro	Arg	Arg	Thr	Gly	Ala	Glu	Val	Asp	Ile	Thr	
				175					180					185			
30	GGC	ATG	ACA	ATG	CTG	CTA	CAA	AAT	CTG	GGG	TAC	AGC	GTA	GAT	GTG	AAA	626
	Gly	Met	Thr	Met	Leu	Leu	Gln	Asn	Leu	Gly	Tyr	Ser	Val	Asp	Val	Lys	
				190				195					200				
35	AAA	AAT	CTC	ACT	GCT	TCG	GAC	ATG	ACT	ACA	GAG	CTG	GAG	GCA	TTT	GCA	674
	Lys	Asn	Leu	Thr	Ala	Ser	Asp	Met	Thr	Thr	Glu	Leu	Glu	Ala	Phe	Ala	
		205					210				215						
	CAC	CGC	CCA	GAG	CAC	AAG	ACC	TCT	GAC	AGC	ACG	TTC	CTG	GTG	TTC	ATG	722
	His	Arg	Pro	Glu	His	Lys	Thr	Ser	Asp	Ser	Thr	Phe	Leu	Val	Phe	Met	
	220					225					230					235	

	TCT CAT GGT ATT CGG GAA GGC ATT TGT GGG AAG AAA CAC TCT GAG CAA	770
	Ser His Gly Ile Arg Glu Gly Ile Cys Gly Lys Lys His Ser Glu Gln	
	240 245 250	
5	GTC CCA GAT ATA CTA CAA CTC AAT GCA ATC TTT AAC ATG TTG AAT ACC	818
	Val Pro Asp Ile Leu Gln Leu Asn Ala Ile Phe Asn Met Leu Asn Thr	
	255 260 265	
	AAG AAC TGC CCA AGT TTG AAG GAC AAA CCG AAG GTG ATC ATC ATC CAG	866
	Lys Asn Cys Pro Ser Leu Lys Asp Lys Pro Lys Val Ile Ile Ile Gln	
	270 275 280	
10	GCC TGC CGT GGT GAC AGC CCT GGT GTG GTG TGG TTT AAA GAT TCA GTA	914
	Ala Cys Arg Gly Asp Ser Pro Gly Val Val Trp Phe Lys Asp Ser Val	
	285 290 295	
	GGA GTT TCT GGA AAC CTA TCT TTA CCA ACT ACA GAA GAG TTT GAG GAT	962
	Gly Val Ser Gly Asn Leu Ser Leu Pro Thr Thr Glu Glu Phe Glu Asp	
15	300 305 310 315	
	GAT GCT ATT AAG AAA GCC CAC ATA GAG AAG GAT TTT ATC GCT TTC TGC	1010
	Asp Ala Ile Lys Lys Ala His Ile Glu Lys Asp Phe Ile Ala Phe Cys	
	320 325 330	
	TCT TCC ACA CCA GAT AAT GTT TCT TGG AGA CAT CCC ACA ATG GGC TCT	1058
20	Ser Ser Thr Pro Asp Asn Val Ser Trp Arg His Pro Thr Met Gly Ser	
	335 340 345	
	GTT TTT ATT GGA AGA CTC ATT GAA CAT ATG CAA GAA TAT GCC TGT TCC	1106
	Val Phe Ile Gly Arg Leu Ile Glu His Met Gln Glu Tyr Ala Cys Ser	
	350 355 360	
25	TGT GAT GTG GAG GAA ATT TTC CGC AAG GTT CGA TTT TCA TTT GAG CAG	1154
	Cys Asp Val Glu Glu Ile Phe Arg Lys Val Arg Phe Ser Phe Glu Gln	
	365 370 375	
	CCA GAT GGT AGA GCG CAG ATG CCC ACC ACT GAA AGA GTG ACT TTG ACA	1202
	Pro Asp Gly Arg Ala Gln Met Pro Thr Thr Glu Arg Val Thr Leu Thr	
30	380 385 390 395	
	AGA TGT TTC TAC CTC TTC CCA GGA CAT TAAAATAAGG AAAGTGTATG AATGTCT	1256
	Arg Cys Phe Tyr Leu Phe Pro Gly His	
	400	
35	GCGGGCAGGA AGTGAAGAGA TCGTTCTGTA AAAGGTTTTT GGAATTATGT CTGCTGAATA	1316
	ATAAACTTTT TTTGAAATAA TAAATCTGGT AGAAAAATGA AAAAAAAAAA AAAAAA	1373

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 404 amino acids

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

	Met	Ala	Asp	Lys	Val	Leu	Lys	Glu	Lys	Arg	Lys	Leu	Phe	Ile	Arg	Ser
	1				5					10					15	
10	Met	Gly	Glu	Gly	Thr	Ile	Asn	Gly	Leu	Xaa	Asp	Glu	Leu	Leu	Gln	Thr
				20					25					30		
	Arg	Val	Leu	Asn	Lys	Glu	Glu	Met	Glu	Lys	Val	Lys	Arg	Glu	Asn	Ala
				35				40					45			
	Thr	Val	Met	Asp	Lys	Thr	Arg	Ala	Leu	Ile	Asp	Ser	Val	Ile	Pro	Lys
				50			55					60				
15	Xaa	Ala	Gln	Ala	Cys	Gln	Ile	Cys	Ile	Thr	Tyr	Ile	Cys	Glu	Glu	Asp
	65					70					75				80	
	Ser	Tyr	Leu	Ala	Gly	Thr	Leu	Gly	Leu	Ser	Ala	Asp	Gln	Thr	Ser	Gly
					85					90					95	
20	Asn	Tyr	Leu	Asn	Met	Gln	Asp	Ser	Gln	Gly	Val	Leu	Ser	Ser	Phe	Pro
				100					105						110	
	Ala	Pro	Gln	Ala	Val	Gln	Asp	Asn	Pro	Ala	Met	Pro	Thr	Ser	Ser	Gly
				115				120					125			
	Ser	Glu	Gly	Asn	Val	Lys	Leu	Cys	Ser	Leu	Glu	Glu	Ala	Gln	Arg	Ile
							135					140				
25	Trp	Lys	Gln	Lys	Ser	Ala	Glu	Ile	Tyr	Pro	Ile	Met	Asp	Lys	Ser	Ser
	145					150					155				160	
	Arg	Thr	Arg	Leu	Ala	Leu	Ile	Ile	Cys	Asn	Glu	Glu	Phe	Asp	Ser	Ile
					165					170					175	
30	Pro	Arg	Arg	Thr	Gly	Ala	Glu	Val	Asp	Ile	Thr	Gly	Met	Thr	Met	Leu
				180					185					190		
	Leu	Gln	Asn	Leu	Gly	Tyr	Ser	Val	Asp	Val	Lys	Lys	Asn	Leu	Thr	Ala
			195					200					205			
	Ser	Asp	Met	Thr	Thr	Glu	Leu	Glu	Ala	Phe	Ala	His	Arg	Pro	Glu	His
		210					215					220				
35	Lys	Thr	Ser	Asp	Ser	Thr	Phe	Leu	Val	Phe	Met	Ser	His	Gly	Ile	Arg
	225					230					235				240	
	Glu	Gly	Ile	Cys	Gly	Lys	Lys	His	Ser	Glu	Gln	Val	Pro	Asp	Ile	Leu
					245					250					255	
40	Gln	Leu	Asn	Ala	Ile	Phe	Asn	Met	Leu	Asn	Thr	Lys	Asn	Cys	Pro	Ser
				260					265					270		
	Leu	Lys	Asp	Lys	Pro	Lys	Val	Ile	Ile	Ile	Gln	Ala	Xaa	Arg	Xaa	Asp
			275					280					285			
	Ser	Pro	Gly	Val	Val	Trp	Phe	Lys	Asp	Ser	Val	Gly	Val	Ser	Gly	Asn
		290					295					300				
45	Leu	Ser	Leu	Pro	Thr	Thr	Glu	Glu	Phe	Glu	Asp	Asp	Ala	Ile	Lys	Lys
	305					310					315				320	
	Ala	His	Ile	Xaa	Lys	Asp	Phe	Ile	Ala	Phe	Cys	Ser	Ser	Thr	Pro	Asp
					325					330					335	

	Asn	Val	Ser	Xaa	Arg	His	Pro	Thr	Met	Gly	Ser	Val	Phe	Ile	Gly	Arg
				340					345					350		
	Leu	Ile	Glu	His	Met	Gln	Glu	Tyr	Xaa	Cys	Ser	Cys	Asp	Val	Glu	Glu
			355					360					365			
5	Ile	Phe	Arg	Lys	Val	Arg	Phe	Ser	Phe	Glu	Gln	Pro	Asp	Gly	Arg	Ala
		370				375						380				
	Gln	Met	Pro	Thr	Thr	Xaa	Arg	Val	Xaa	Leu	Thr	Arg	Cys	Phe	Tyr	Leu
	385					390					395					400
	Phe	Pro	Gly	His												

10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 505 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

20	Met	Met	Arg	Gln	Asp	Arg	Trp	Leu	Leu	Glu	Arg	Asn	Ile	Leu	Glu	Phe
	1				5					10					15	
	Ser	Ser	Lys	Leu	Gln	Ala	Asp	Leu	Ile	Leu	Asp	Val	Leu	Ile	Ala	Lys
				20					25					30		
	Gln	Val	Leu	Asn	Ser	Asp	Asn	Gly	Asp	Val	Ile	Asn	Ser	Cys	Arg	Thr
			35					40					45			
25	Glu	Arg	Asp	Asn	Glu	Lys	Glu	Ile	Val	Lys	Ala	Val	Gln	Arg	Arg	Gly
		50					55					60				
	Asp	Glu	Ala	Phe	Asp	Ala	Phe	Tyr	Asp	Ala	Leu	Arg	Asp	Thr	Gly	His
	65					70					75					80
	Asn	Asp	Leu	Ala	Asp	Val	Leu	Met	Pro	Leu	Ser	Arg	Pro	Xaa	Xaa	Xaa
30					85					90					95	
	Asn	Pro	Val	Pro	Met	Glu	Cys	Pro	Met	Ser	Pro	Ser	Ser	His	Arg	Arg
				100					105					110		
	Ser	Arg	Ala	Leu	Ser	Pro	Pro	Xaa	Tyr	Ala	Ser	Pro	Thr	Arg	Val	His
			115					120					125			
35	Arg	Asp	Ser	Ile	Ser	Ser	Val	Ser	Ser	Phe	Thr	Ser	Thr	Tyr	Gln	Asp
		130					135					140				
	Val	Tyr	Ser	Arg	Ala	Arg	Ser	Ser	Ser	Arg	Ser	Ser	Arg	Pro	Leu	Gln
	145					150				155					160	
	Ser	Ser	Asp	Arg	His	Asn	Tyr	Met	Ser	Ala	Ala	Thr	Ser	Phe	Pro	Ser
40					165					170					175	
	Gln	Pro	Xaa	Ser	Ala	Asn	Ser	Ser	Phe	Thr	Gly	Cys	Ala	Ser	Leu	Gly
				180					185					190		
	Tyr	Ser	Ser	Ser	Arg	Asn	Arg	Ser	Phe	Ser	Lys	Thr	Ser	Ala	Gln	Ser
			195					200					205			
45	Gln	Tyr	Ile	Phe	His	Glu	Glu	Asp	Met	Asn	Tyr	Val	Asp	Ala	Pro	Thr
		210					215					220				

	Ile	His	Arg	Val	Phe	Asp	Glu	Lys	Thr	Met	Tyr	Arg	Asn	Phe	Ser	Ser
	225					230					235					240
	Pro	Arg	Gly	Leu	Cys	Leu	Ile	Ile	Asn	Asn	Glu	His	Phe	Glu	Gln	Met
					245					250						255
5	Pro	Thr	Arg	Asn	Gly	Thr	Lys	Ala	Asp	Lys	Asp	Asn	Leu	Thr	Asn	Ile
				260					265					270		
	Phe	Arg	Cys	Met	Gly	Tyr	Thr	Val	Ile	Cys	Lys	Asp	Asn	Leu	Thr	Gly
			275					280					285			
10	Arg	Glu	Met	Leu	Ser	Thr	Ile	Arg	Ser	Phe	Gly	Arg	Asn	Asp	Met	His
		290					295					300				
	Gly	Asp	Ser	Ala	Ile	Leu	Val	Ile	Leu	Ser	His	Gly	Glu	Xaa	Asn	Val
						310					315					320
	Ile	Ile	Gly	Val	Asp	Asp	Val	Ser	Val	Asn	Val	His	Glu	Ile	Tyr	Asp
					325					330						335
15	Leu	Leu	Asn	Ala	Ala	Asn	Ala	Pro	Arg	Leu	Ala	Asn	Lys	Pro	Lys	Leu
				340					345					350		
	Val	Phe	Val	Gln	Ala	Cys	Arg	Gly	Glu	Arg	Arg	Asp	Asn	Gly	Phe	Pro
			355					360					365			
20	Val	Leu	Asp	Ser	Val	Asp	Gly	Val	Pro	Ser	Leu	Ile	Arg	Arg	Gly	Trp
		370					375					380				
	Asp	Asn	Arg	Asp	Gly	Pro	Leu	Phe	Asn	Phe	Leu	Gly	Cys	Val	Arg	Pro
						390					395					400
	Gln	Val	Gln	Gln	Val	Trp	Arg	Lys	Lys	Pro	Ser	Gln	Ala	Asp	Met	Leu
					405					410					415	
25	Ile	Ala	Tyr	Ala	Thr	Thr	Ala	Gln	Tyr	Val	Ser	Trp	Arg	Asn	Ser	Ala
				420					425					430		
	Arg	Gly	Ser	Trp	Phe	Ile	Gln	Ala	Val	Cys	Glu	Val	Phe	Ser	Leu	His
			435					440					445			
30	Ala	Lys	Asp	Met	Asp	Val	Val	Glu	Leu	Leu	Thr	Glu	Val	Asn	Lys	Lys
		450					455					460				
	Val	Ala	Cys	Gly	Phe	Gln	Thr	Ser	Gln	Gly	Ser	Asn	Ile	Leu	Lys	Gln
		465				470					475					480
	Met	Pro	Glu	Leu	Thr	Ser	Arg	Leu	Leu	Lys	Lys	Phe	Tyr	Phe	Trp	Pro
					485					490					495	
35	Glu	Asp	Arg	Gly	Arg	Asn	Ser	Ala	Val							
				500					505							

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 479 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

45 Thr Val Ser Leu Ser Leu Ile Ile Ala Arg Gln Val Leu Asn Ser Asp
1 5 10 15

	Asn	Gly	Asp	Met	Ile	Asn	Ser	Cys	Arg	Thr	Glu	Arg	Asp	Asn	Glu	Lys
				20					25					30		
	Glu	Ile	Val	Lys	Ala	Val	Gln	Arg	Arg	Gly	Asp	Glu	Ala	Phe	Asp	Ala
			35					40					45			
5	Phe	Tyr	Asp	Ala	Leu	Arg	Asp	Thr	Gly	His	Asn	Asp	Leu	Ala	Asp	Val
		50					55					60				
	Leu	Met	Pro	Leu	Ser	Arg	Pro	Val	Asp	Ser	Asn	Pro	Val	Pro	Met	Glu
	65					70					75					80
	Cys	Pro	Met	Ser	Pro	Ser	Ser	His	Arg	Arg	Ser	Arg	Ala	Leu	Ser	Pro
10					85					90					95	
	Pro	Xaa	Tyr	Ala	Ser	Pro	Thr	Arg	Val	His	Arg	Asp	Ser	Ile	Ser	Ser
				100					105					110		
	Val	Ser	Ser	Phe	Thr	Ser	Thr	Tyr	Gln	Asp	Val	Tyr	Ser	Arg	Ala	Thr
			115					120					125			
15	Ser	Ser	Ser	Pro	Leu	Gln	Thr	Ser	Asp	Arg	His	Asn	Tyr	Val	Ser	Ala
		130					135					140				
	Ser	Thr	Ser	Phe	Gln	Ser	Gln	Pro	Ala	Ser	Ala	Asn	Ser	Ser	Phe	Thr
	145					150					155					160
	Gly	Ser	Ala	Ser	Leu	Gly	Tyr	Ser	Ser	Ser	Arg	Thr	Arg	Ser	Tyr	Ser
20					165					170					175	
	Lys	Thr	Ser	Ala	His	Ser	Gln	Tyr	Ile	Phe	His	Glu	Glu	Asp	Met	Asn
				180					185					190		
	Tyr	Val	Asp	Ala	Pro	Thr	Ile	His	Arg	Val	Phe	Asp	Glu	Lys	Thr	Met
			195					200					205			
25	Tyr	Arg	Asn	Phe	Ser	Thr	Pro	Arg	Gly	Leu	Cys	Leu	Ile	Ile	Asn	Asn
		210					215					220				
	Glu	His	Phe	Glu	Gln	Met	Pro	Thr	Arg	Asn	Gly	Thr	Lys	Pro	Asp	Lys
	225					230					235					240
	Asp	Asn	Ile	Ser	Asn	Ile	Phe	Arg	Cys	Met	Gly	Tyr	Ile	Val	His	Cys
30					245					250					255	
	Lys	Asp	Asn	Leu	Thr	Gly	Arg	Glu	Met	Met	Ser	Thr	Ile	Arg	Ser	Phe
			260					265					270			
	Gly	Arg	Asn	Asp	Thr	His	Gly	Asp	Ser	Ala	Ile	Leu	Val	Ile	Leu	Ser
			275				280					285				
35	His	Gly	Glu	Xaa	Asn	Val	Ile	Ile	Gly	Val	Asp	Asp	Val	Ser	Val	Asn
		290				295						300				
	Val	His	Glu	Ile	Tyr	Xaa	Leu	Leu	Asn	Ala	Ala	Asn	Ala	Pro	Arg	Leu
	305					310					315					320
	Ala	Asn	Lys	Pro	Lys	Leu	Val	Phe	Val	Gln	Ala	Cys	Arg	Gly	Glu	Arg
40					325					330					335	
	Arg	Asp	Val	Gly	Phe	Pro	Val	Leu	Asp	Ser	Val	Asp	Gly	Val	Pro	Ser
			340						345				350			
	Leu	Ile	Arg	Arg	Gly	Trp	Asp	Lys	Gly	Asp	Gly	Pro	Leu	Phe	Asn	Phe
		355					360					365				
45	Leu	Gly	Cys	Val	Arg	Pro	Gln	Ala	Gln	Gln	Val	Trp	Arg	Lys	Lys	Pro
		370				375					380					
	Ser	Gln	Ala	Asp	Met	Leu	Ile	Ala	Tyr	Ala	Thr	Thr	Ala	Gln	Tyr	Val
	385					390					395					400
	Ser	Trp	Arg	Asn	Ser	Ala	Arg	Gly	Ser	Trp	Phe	Ile	Gln	Ala	Val	Cys
50					405					410					415	
	Glu	Val	Phe	Ser	Leu	His	Ala	Lys	Asp	Met	Asp	Val	Val	Glu	Leu	Leu

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

5 GGTTTTAACC AGTTACTCAA G

21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCGGTGACAT TGGACACTC

19

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACTATTCAAC ACTTG

15

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 171 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (D) OTHER INFORMATION: Xaa at position 117 is Ala or Val.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Leu Thr Val Gln Val Tyr Arg Thr Ser Gln Lys Cys Ser Ser Ser
1 5 10 15

[illegible]

35 (2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: Xaa at position 3 is Ala, His, Gln, Lys, Phe, Cha or Asp.

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Tyr Val Xaa Asp

1

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: Peptide has an acetyl group at the amino terminus and an aldehyde group at the carboxyl terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Tyr Val Ala Asp

1

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Tyr Val Ala Asp

1

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: Xaa at position 2 is D-Ala. Peptide has an acetyl group at the amino terminus and an aldehyde group at the carboxyl terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Tyr Xaa Ala Asp

1

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: Peptide has an acetyl group at the amino terminus and an aldehyde group at the carboxyl terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Tyr Val Lys Asp

1

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 354 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: C-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Arg	Ser	Arg	Ser	Arg	Ser	Arg	Ala	Leu	His	Ser	Ser	Asp	Arg	His	Asn
1				5					10					15	
Tyr	Ser	Ser	Pro	Pro	Val	Asn	Ala	Phe	Pro	Ser	Gln	Pro	Ser	Ser	Ala
			20					25					30		
Asn	Ser	Ser	Phe	Thr	Gly	Cys	Ser	Ser	Leu	Gly	Tyr	Ser	Ser	Ser	Arg
			35				40					45			
Asn	Arg	Ser	Phe	Ser	Lys	Ala	Ser	Gly	Pro	Thr	Gln	Tyr	Ile	Phe	His
			50				55				60				
Glu	Glu	Asp	Met	Asn	Phe	Val	Asp	Ala	Pro	Thr	Ile	Ser	Arg	Val	Phe
65				70					75					80	
Asp	Glu	Lys	Thr	Met	Tyr	Arg	Asn	Phe	Ser	Ser	Pro	Arg	Gly	Met	Cys
				85					90					95	
Leu	Ile	Ile	Asn	Asn	Glu	His	Phe	Glu	Gln	Met	Pro	Thr	Arg	Asn	Gly
			100					105					110		
Thr	Lys	Ala	Asp	Lys	Asp	Asn	Leu	Thr	Asn	Leu	Phe	Arg	Cys	Met	Gly
			115				120					125			
Tyr	Thr	Val	Ile	Cys	Lys	Asp	Asn	Leu	Thr	Gly	Arg	Gly	Met	Leu	Leu

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 294 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Phe	Pro	Ala	Pro	Gln	Ala	Val	Gln	Asp	Asn	Pro	Ala	Met	Pro	Thr	Ser
1				5				10						15	
Ser	Gly	Ser	Glu	Gly	Asn	Val	Lys	Leu	Cys	Ser	Leu	Glu	Glu	Ala	Gln
			20				25					30			
Arg	Ile	Trp	Lys	Gln	Lys	Ser	Ala	Glu	Ile	Tyr	Pro	Ile	Met	Asp	Lys
		35				40					45				
Ser	Ser	Arg	Thr	Arg	Leu	Ala	Leu	Ile	Ile	Cys	Asn	Glu	Glu	Phe	Asp
	50				55					60					
Ser	Ile	Pro	Arg	Arg	Thr	Gly	Ala	Glu	Val	Asp	Ile	Thr	Gly	Met	Thr
65					70					75				80	
Met	Leu	Leu	Gln	Asn	Leu	Gly	Tyr	Ser	Val	Asp	Val	Lys	Lys	Asn	Leu
			85					90						95	
Thr	Ala	Ser	Asp	Met	Thr	Thr	Glu	Leu	Glu	Ala	Phe	Ala	His	Arg	Pro
			100					105					110		
Glu	His	Lys	Thr	Ser	Asp	Ser	Thr	Phe	Leu	Val	Phe	Met	Ser	His	Gly
		115					120					125			
Ile	Arg	Glu	Gly	Ile	Cys	Gly	Lys	Lys	His	Ser	Glu	Gln	Val	Pro	Asp
	130					135					140				
Ile	Leu	Gln	Leu	Asn	Ala	Ile	Phe	Asn	Met	Leu	Asn	Thr	Lys	Asn	Cys
145					150					155				160	
Pro	Ser	Leu	Lys	Asp	Lys	Pro	Lys	Val	Ile	Ile	Ile	Gln	Ala	Xaa	Arg
			165					170						175	
Xaa	Asp	Ser	Pro	Gly	Val	Val	Trp	Phe	Lys	Asp	Ser	Val	Gly	Val	Ser
		180						185					190		
Gly	Asn	Leu	Ser	Leu	Pro	Thr	Thr	Glu	Glu	Phe	Glu	Asp	Asp	Ala	Ile
		195					200					205			
Lys	Lys	Ala	His	Ile	Xaa	Lys	Asp	Phe	Ile	Ala	Phe	Cys	Ser	Ser	Thr
	210					215						220			
Pro	Asp	Asn	Val	Ser	Xaa	Arg	His	Pro	Thr	Met	Gly	Ser	Val	Phe	Ile
225					230					235				240	
Gly	Arg	Leu	Ile	Glu	His	Met	Gln	Glu	Tyr	Xaa	Cys	Ser	Cys	Asp	Val
			245					250						255	
Glu	Glu	Ile	Phe	Arg	Lys	Val	Arg	Phe	Ser	Phe	Glu	Gln	Pro	Asp	Gly
		260						265					270		
Arg	Ala	Gln	Met	Pro	Thr	Thr	Xaa	Arg	Val	Xaa	Leu	Thr	Arg	Cys	Phe
		275					280						285		

Tyr Leu Phe Pro Gly His
290

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ser Val Gly Val Ser Gly Asn Leu Ser Leu Pro Thr Thr Glu Glu Phe
1 5 10 15
Glu Asp Asp Ala Ile Lys Lys Ala His Ile Xaa Lys Asp Phe Ile Ala
20 25 30
Phe Cys Ser Ser Thr Pro Asp Asn Val Ser Xaa Arg His Pro Thr Met
35 40 45
Gly Ser Val Phe Ile Gly Arg Leu Ile Glu His Met Gln Glu Tyr Xaa
50 55 60
Cys Ser Cys Asp Val Glu Glu Ile Phe Arg Lys Val Arg Phe Ser Phe
65 70 75 80
Glu Gln Pro Asp Gly Arg Ala Gln Met Pro Thr Thr Xaa Arg Val Xaa
85 90 95
Leu Thr Arg Cys Phe Tyr Leu Phe Pro Gly His
100 105

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 187 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Phe Pro Ala Pro Gln Ala Val Gln Asp Asn Pro Ala Met Pro Thr Ser
1 5 10 15
Ser Gly Ser Glu Gly Asn Val Lys Leu Cys Ser Leu Glu Glu Ala Gln
20 25 30
Arg Ile Trp Lys Gln Lys Ser Ala Glu Ile Tyr Pro Ile Met Asp Lys
35 40 45
Ser Ser Arg Thr Arg Leu Ala Leu Ile Ile Cys Asn Glu Glu Phe Asp

145 150 155 160
 Gln Gln Leu Tyr Leu Phe Pro Gly Tyr Pro Pro Thr
 165 170

(2) INFORMATION FOR SEQ ID NO:27:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 172 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: protein
 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

15 Met Leu Thr Val Gln Val Tyr Arg Thr Ser Gln Lys Cys Ser Ser Ser
 1 5 10 15
 20 Lys His Val Val Glu Val Leu Leu Asp Pro Leu Gly Thr Ser Phe Cys
 20 25 30
 Ser Leu Leu Pro Pro Pro Leu Leu Leu Tyr Glu Thr Asp Arg Gly Val
 35 40 45
 25 Asp Gln Gln Asp Gly Lys Asn His Thr Gln Ser Pro Gly Cys Glu Glu
 50 55 60
 Ser Asp Ala Gly Lys Glu Glu Leu Met Lys Met Arg Leu Pro Thr Arg
 65 70 75 80
 Ser Asp Met Ile Cys Gly Tyr Ala Cys Leu Lys Gly Asn Ala Ala Met
 85 90 95
 30 Arg Asn Thr Lys Arg Gly Ser Trp Tyr Ile Glu Ala Leu Thr Gln Val
 100 105 110
 Phe Ser Glu Arg Val Xaa Cys Asp Met His Val Ala Asp Met Leu Val
 115 120 125
 Lys Val Asn Ala Leu Ile Lys Glu Arg Glu Gly Tyr Ala Pro Gly Thr
 130 135 140
 Glu Phe His Arg Cys Lys Glu Met Ser Glu Tyr Cys Ser Thr Leu Cys
 145 150 155 160
 Gln Gln Leu Tyr Leu Phe Pro Gly Tyr Pro Pro Thr
 165 170